

(19)



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11)

EP 0 647 710 B1

(12)

EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention
of the grant of the patent:
20.03.2002 Bulletin 2002/12

(51) Int Cl.⁷: **C12N 9/52, C12N 15/57**
// (C12N9/52, C12R1:465)

(21) Application number: **93911969.9**

(86) International application number:
PCT/JP93/00592

(22) Date of filing: **30.04.1993**

(87) International publication number:
WO 93/23530 (25.11.1993 Gazette 1993/28)

(54) **NOVEL PROTEASE**

NEUARTIGE PROTEASE

NOUVELLE PROTEASE

(84) Designated Contracting States:
**AT BE CH DE DK ES FR GB GR IE IT LI LU MC NL
PT SE**

(30) Priority: **19.05.1992 JP 12651192**

(43) Date of publication of application:
12.04.1995 Bulletin 1995/15

(73) Proprietor: **SHIONOGI SEIYAKU KABUSHIKI
KAISHA**
Osaka 541 (JP)

(72) Inventors:
• **NAKAMURA, Etsuo 4573-41, Karato Arino-cho
Kita-ku
Hyogo 651-13 (JP)**
• **TSUZUKI, Hiroshige 5-12-15, Osumigaoka
Tanabe-cho
Kyoto 610-03 (JP)**
• **KITADOKORO, Kengo 7-47-202, Minoo 6-chome
Osaka 562 (JP)**
• **SHIN, Masaru 3-110, Koryo-cho Kita-ku
Hyogo 651-12 (JP)**
• **TERAOKA, Hiroshi 2-47-10, Takakuradai
Osaka 590-01 (JP)**

(74) Representative: **Nachshen, Neil Jacob et al
D Young & Co
21 New Fetter Lane
London EC4A 1DA (GB)**

(56) References cited:

- **INT J BIOCHEM 23 (10), 1991, 979-984, CODEN: IJBOV ISSN: 0020-711X, XP000568704 SINHA U ET AL: "TWO NEW EXTRACELLULAR SERINE PROTEASES FROM STREPTOMYCES - FRADIAE."**
- **BIOCHIM BIOPHYS ACTA 1163 (2), 1993, 149-157, CODEN: BBACAQ ISSN: 0006-3002, XP000568705 KITADOKORO K ET AL: "PURIFICATION CHARACTERIZATION AND MOLECULAR CLONING OF AN ACIDIC AMINO ACID-SPECIFIC PROTEINASE FROM STREPTOMYCES - FRADIAE ATCC 14544."**
- **European Journal of Biochemistry, Vol. 204, No. 1 (1992), I. SVENDSEN et al., "Isolation and Amino Acid Sequence of a Glutamic Acid Specific Endopeptidase from Bacillus licheniformis", p. 165-171.**
- **FEBS (Federation of European Biochemical Societies), Vol. 292, Nos. 1-2 (1991), I SVENDSEN et al., "The Primary Structure of the Glutamic Acid-Specific Protease of Streptomyces griseus", p. 165-167.**
- **Journal of Biochemistry, Vol. 104, No. 3 (1988), (Tokyo), N. YOSHIDA et al., "Purification and Characterization of an Acidic Amino Acid Specific Endopeptidase of Streptomyces griseus Obtained from a Commercial Preparation (Pronase)", p. 451-456.**

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

EP 0 647 710 B1

DescriptionField of the Invention

[0001] The present invention relates to a novel protease for cleaving the carboxyl side of a glutamic acid residue in an amino acid sequence of a polypeptide, and a DNA sequence for encoding the protease.

Background of the Invention

[0002] As an enzyme for acting on a protein (polypeptide) so as to specifically cleave the carboxyl side of a glutamic acid (Glu) residue, V8 protease derived from *Staphylococcus aureus* V8 strain is known (Gabriel R. Drapeau *et al.*, J. Biol. Chem. 247, 20, 6720-6726 (1972)). This enzyme is classified as a serine protease. C. Carmona *et al.* (Nucl. Acids Res., 15, 6757 (1987)) cloned a DNA sequence for encoding this enzyme.

[0003] A similar enzyme, acidic amino acid specific endopeptidase derived from *Streptomyces griseus*, which is an actinomycetes, is also known (Norio Yoshida *et al.*, J. Biochem. 104, 3, 451-456 (1988)). Recently, I Svendsen *et al.* reported an amino acid sequence of the acidic amino acid specific endopeptidase purified from *Streptomyces griseus* (FEBS LETTERS 292 165-167 (1991)). Furthermore, glutamic acid specific endopeptidase derived from *Bacillus subtilis* is also known (Takuro Niidome *et al.*, J. Biochem. 108, 965-970 (1990)). With respect to some of these proteases, the characteristics thereof have been studied and inhibitors against them are known (K. Nagata *et al.*, J. Biochem. 110, 859-862 (1991) and T. Komiyama, J. Biol. Chem. 266, 17, 10727-10730 (1991)).

[0004] The aforementioned enzymes are useful for specifically cleaving a protein at the aforementioned site for the purpose of structural analysis of the protein; for obtaining an objective protein through cleavage of a fusion protein in the case where the objective protein is desired to be produced as the fusion protein by a genetic recombination technique and the like. In the latter case, for example, after producing the objective protein by coupling with another protein through a Glu residue, the resultant is subjected to cleavage with one of the enzymes.

[0005] Thus, the objective protein can be separated. Accordingly, there is a demand for a protease having such an enzymatic activity other than the above-mentioned enzymes.

Disclosure of the Invention:

[0006] The objective of this invention is producing a novel polypeptide having protease activity for cleaving the carboxyl side of a Glu residue, and a DNA sequence for encoding the protease.

[0007] The present inventors have made various studies to obtain a polypeptide having the activity of an acidic amino acid specific endopeptidase (in particular, a glutamic acid specific endopeptidase) from a strain of a microorganism other than the above-mentioned microorganisms. As a result, they found a novel polypeptide having the above-mentioned microorganisms. As a result, they found a novel polypeptide having the above-mentioned characteristic derived from *Streptomyces fradiae*, and studied the characteristics thereof in detail. Furthermore, they determined a DNA sequence for encoding this polypeptide thereby accomplishing the present invention.

[0008] The polypeptide of this invention comprises a mature peptide and a signal peptide, wherein said mature peptide is characterized by having protease activity for cleaving a peptide bond including a carboxyl group of a glutamic acid residue in a peptide and having an amino acid sequence from Val in the 1 position to Tyr in the 187 position of SEQ ID No. 1 and wherein said signal peptide comprises an amino acid sequence from Met in the -170 position to Tyr in the 187 position of SEQ ID No. 1.

[0009] In a preferred embodiment, the protease is derived from *Streptomyces fradiae*.

[0010] In a preferred embodiment, the protease is derived from *Streptomyces fradiae* ATCC 14544 strain. The DNA sequence of this invention encodes any one of the aforementioned polypeptides.

[0011] In a preferred embodiment, the DNA sequence comprises a base sequence from A in the 435 position to C in the 1505 position of SEQ ID No. 1.

[0012] The polypeptide of the invention (hereinafter referred to as the SFase) is produced from *Streptomyces fradiae*. For example, it is produced from *Streptomyces fradiae* ATCC 14544 strain available from American Type Culture Collection (ATCC).

Brief Description of the Drawings:

[0013] Figure 1 shows the comparison in the amino acid sequences between a mature peptide of SFase of the invention (upper lines) and a mature peptide of the same kind of an enzyme derived from *Streptomyces griseus* (lower lines). The portions shown with dots indicate the same amino acids.

[0014] Figure 2 shows an elution pattern of the SFase from S-Sepharose.

Preferred Embodiments for Practicing the Invention:

[0015] The present invention will now be described in detail in the order of the procedures.

I. Culture condition:

[0016] The SFase of the invention is produced by using Streptomyces fradiae, for example, Streptomyces fradiae ATCC 14544 strain. In the case of culturing the strain that can produce the SFase, a medium used for the strain is not required to be special, and any ordinary medium can be used. For example, a medium including glucose, soybean meal, yeast extract, corn steep liquor, potato starch and various salts may be used. The medium pH is 5 through 9, and preferably approximately 7.0, and the medium temperature is 15°C through 50°C, and preferably approximately 28°C. The strain is cultured, for example, aerobically for approximately 3 through 5 days with stirring or shaking. The SFase is mainly secreted extracellularly.

II. Method for harvesting the enzyme:

[0017] In order to harvest and purify the SFase of the invention from the culture, one or a combination of known production methods can be used. For example, the culture solution may be centrifuged to obtain a supernatant, which may be purified by an appropriate means, thereby obtaining the enzyme of the invention.

[0018] For example, the supernatant may be precipitated with ammonium sulfate. The precipitate may be dissolved in a buffer solution, and the mixture may be subjected to separation by chromatography using S-Sepharose. Then, the resultant may be subjected to affinity chromatography using S-Sepharose again, thereby giving the present enzyme. The characteristics of the enzyme described below was measured by using this enzyme preparation.

III. Method for measuring the activity:

[0019] As a substrate, Z-Phe-Leu-Glu-pNA (wherein Z indicates carbobenzoxy), and pNA indicates a p-nitroaniline residue) is used, and the substrate is dissolved in 50 mM Tris-HCl (pH 7.5, including 2 mM CaCl_2 and 2% DMF) so as to achieve a final concentration of 0.2 mM. An enzyme solution is added thereto, and allowed to react at a temperature of 37°C for 10 minutes. The absorbance of p-nitroaniline released through the enzymatic reaction into the reaction solution is measured at 410 nm. The amount of the enzyme for attaining an absorbance of 1.0 is defined as 1 unit (U).

IV. Characteristics of the enzyme:

[0020] The enzymatic and protein chemical characteristics of the SFase of the invention are as follows:

(1) Enzymatic action and substrate specificity:

[0021]

1. Synthetic substrates listed in Table 1 were prepared and respectively dissolved in 50 mM Tris-HCl (including 2 mM CaCl_2 , and dimethylformamide (DMF) or dimethyl sulfoxide (DMSO) in the ratio shown in Table 1; pH 7.5) so as to achieve respective concentrations as listed in Table 1. The present enzyme was added thereto to cause a reaction at a temperature of 25°C. By measuring the absorbance (410 nm) of p-nitroaniline released through the enzymatic reaction into the reaction solution, the amount (A410 nm) of p-nitroaniline released from each substrate per 1 minute was calculated. The results are listed in Table 1.

Table 1:
Substrate specificity:

| Substrate | Concentration (final) mM | Solvent(%) | A 410 nm /mg/min |
|---------------------|-----------------------------|------------|---------------------|
| Z-Phe-Leu-Glu-pNA | 0.20 | DMF 0.2~ | 319.1 |
| Boc-Ala-Ala-Glu-pNA | 0.20 | DMF 2.0 | 126.2 |
| Boc-Ala-Ala-Asp-pNA | 0.20 | DMF 2.0 | 14.3 |
| Bz-Tyr-pNA | 0.02 | DMF 2.0 | 0*) |
| Z-Trp-pNA | 0.20 | DMF 2.0 | 0*) |
| Bz-Arg-pNA | 0.50 | DMSO 1.0 | 0*) |
| Z-Gly-Ser-pNA | 0.20 | DMF 2.0 | 0*) |
| Suc-Ala-Ala-Ala-pNA | 0.20 | - - | 0*) |
| Z-Ala-Ala-Leu-pNA | 0.125 | DMF 5.0 | 0*) |
| Leu-pNA | 2.00 | - - | 0*) |

Condition: 50 mM Tris-HCl, 2 mM CaCl₂,
pH 7.5, 25°C

Detection: A410 nm

DMF: N,N-dimethylformamide

DMSO: dimethyl sulfoxide

Z-: benzyloxycarbonyl-

Boc-: t-butoxycarbonyl-

Bz-: benzoyl-

Suc-: succinyl-

pNA: p-nitroanilide

*) No color detected at A410 nm through a
reaction for 15 minutes at a concentration of the
enzyme of 4.2 µg/ml at a temperature of 23°C.

2. As a proteinous substrate, an oxidized B chain of insulin was selected, and the action of the present enzyme against the substrate was examined. First, the oxidized B chain of insulin was dissolved in 50 mM Tris-HCl, pH 7.5, and the present enzyme was added thereto so as to achieve the enzyme/substrate = 1/139 (W/W), and the reaction was effected for a determined period of time. The obtained reaction mixture was subjected to HPLC using

a Vydac Protein C4 column, and eluted for 30 minutes by a linear concentration gradient using 0% through 50% acetonitrile in 0.1% TFA. The thus obtained fragments were analyzed with regard to the amino acid sequence. As a result, it was found that a bond between Glu13 and Ala14 alone was cleaved.

[0022] From the results of the above described items 1. and 2., it is apparent that the SFase of the invention is a glutamic acid specific endopeptidase for cleaving a peptide bond including carboxyl in a glutamic acid residue in a peptide.

(2) Optimal pH and stable pH range:

[0023] As a substrate, Z-Phe-Leu-Glu-pNA was used. The substrate was dissolved in 50 mM buffers including 10% DMF (an acetate buffer (pH 3.2 through 6.1), a phosphate buffer (pH 5.5 through 8.1), a Tris-HCl buffer (pH 6.1 through 9.8) and a glycinate buffer (pH 9.1 through 11.1)). The SFase of the invention was added thereto, and the reaction was effected at a temperature of 25°C for 10 minutes. The absorbance of p-nitroaniline released through the enzymatic reaction into the reaction solutions was measured at 410 nm. When the reactions were effected in the reaction solutions having various pH, it was found that the optimal pH of the reaction was approximately 8.2.

[0024] Next, under various pH conditions, the SFase of the invention was allowed to stand at a temperature of 37°C for 1.5 hours, and the reactions were effected in accordance with the activity measuring method, thereby measuring the residual activity. As a result, the stable pH range was found to be 6.0 through 9.0.

(3) Effect of an inhibitor:

[0025] The effects of various inhibitors on the SFase of the invention are listed in Table 2. The SFase of the invention is completely inhibited by diisopropyl fluorophosphate (DFP), but is not inhibited by ethylenediaminetetraacetic acid (EDTA) and o-phenanthroline (OP), which are inhibitors of metal proteases. From this, the SFase of the invention was found to be classified as a serine protease.

[0026] Further, the SFase of the invention is inhibited by neither Tosyl-Lys-chloromethylketone (TLCK) nor Tosyl-Phe-chloromethylketone (TPCK), which are inhibitors of trypsin and chymotrypsin, but is completely inhibited by Boc-Phe-Leu-Glu-chloromethylketone. From this, similarly to the result obtained with regard to the substrate specificity, the SFase of the invention is found to be a glutamic acid specific endopeptidase.

Table 2:
Inhibition:

| Inhibitor | Concentration of Inhibitor | Residual Activity (%) |
|--------------------|-------------------------------|--------------------------|
| Control | - | 100 |
| Z-Phe-Leu-Glu-CK | 0.5 mM | 0 |
| Boc-Ala-Ala-Ala-CK | 0.5 mM | 102 |
| TPCK | 0.5 mM | 101 |
| TLCK | 0.5 mM | 112 |
| DFP | 2.0 mM | 7 |
| DFP | 0.5 mM | 20 |
| DFP | 0.25 mM | 33 |
| EDTA | 10.0 mM | 126 |
| EDTA | 1.0 mM | 113 |
| EDTA | 0.1 mM | 110 |
| OP | 10.0 mM | 104 |
| OP | 1.0 mM | 108 |
| OP | 0.1 mM | 106 |

Buffer: 50 mM Tris-HCl, 2.0 mM CaCl₂, pH 7.5,

25°C

Preincubation: 37°C, 1 hour

Enzyme: 1.0 µg/ml

Substrate: 0.2 mM Z-Phe-Leu-Glu-pNA, 0.2% DMF

(4) Molecular weight:

[0027] By SDS-PAGE (15% gel, 1.0 mm; molecular weight marker: RAINBOW™ Protein Molecular Weight Marker manufactured by Amersham), the molecular weight of the present enzyme was measured, and as a result, it was calculated to be 19,000. The molecular weight calculated based on the amino acid sequence determined from the gene sequence of the present enzyme described below is 18,702, which coincides with the value obtained by SDS-PAGE.

(5) Amino acid composition:

[0028] The SFase of the invention was hydrolyzed by using 4 M methanesulfonic acid [including 0.2% 3-(2-aminoethyl)indole] at a temperature of 110°C for determined periods of time (24, 48 and 72 hours). The respective hydrolyzed

products were analyzed with regard to amino acids by a Hitachi amino acid analyzer (Model 835). The results are listed in Table 3. The amino acid composition determined based on the DNA sequence of the present enzyme (described below) are listed as well. These results apparently coincide with each other.

Table 3:
Amino acid composition:

| | Measured Value | Value estimated from cDNA |
|-------|-------------------|---------------------------------|
| Asp | 19.4 | 18 |
| Thr | 21.5 | 21 |
| Ser | 22.9 | 23 |
| Glu | 5.9 | 5 |
| Pro | 2.9 | 3 |
| Gly | 33.5 | 31 |
| Ala | 20.5 | 19 |
| Cys/2 | 4.4 | 4 |
| Val | 18.9 | 20 |
| Met | 1.1 | 1 |
| Ile | 6.1 | 7 |
| Leu | 5.9 | 5 |
| Tyr | 10.5 | 10 |
| Phe | 4.3 | 4 |
| Lys | 4.0 | 4 |
| His | 3.8 | 4 |
| Arg | 7.1 | 7 |
| Trp | 0.9 | 1 |
| Res | - | 187 |
| Mol | - | 18702.4 |

(6) Partial sequence of the amino acid:

1. N-terminal amino acid sequence:

[0029] The amino acid sequence around the N-terminal was analyzed by using a 477A Protein Sequencer manu-

factured by Applied Biosystems. In the analysis, the present enzyme was previously inhibited by DFP. The amino acid sequence from the N-terminal to the 31 residue (corresponding to the amino acids from Val in the 1 position to Ala in the 31 position of SEQ ID No. 1) is shown in Table 4.

Table 4:
N-terminal sequence:

| Step | Amino Acid | pmol | Step | Amino Acid | pmol |
|------|------------|--------|------|------------|--------|
| 1 | Val | 2615.1 | 17 | Ala - | 1048.6 |
| 2 | Ala | 1771.8 | 18 | Phe | 995.6 |
| 3 | Gly | 2175.8 | 19 | Asn | 670.1 |
| 4 | Gly | 2378.4 | 20 | Val | 1080.1 |
| 5 | Asp | 1097.2 | 21 | Thr | 660.4 |
| 6 | Ala | 2195.4 | 22 | Lys | 410.8 |
| 7 | Ile | 2180.3 | 23 | Asn | 455.1 |
| 8 | Tyr | 1341.4 | 24 | Gly | 979.0 |
| 9 | Gly | 1014.5 | 25 | Val | 1054.4 |
| 10 | Gly | 2154.5 | 26 | Arg | 647.2 |
| 11 | Gly | 2227.4 | 27 | Tyr | 464.9 |
| 12 | Ser | 430.6 | 28 | Phe | 396.5 |
| 13 | Arg | 304.6 | 29 | Leu | 514.0 |
| 14 | x | x | 30 | Thr | 280.3 |
| 15 | Ser | 443.0 | 31 | Ala | 580.5 |
| 16 | Ala | 1002.9 | | | |

| | | | |
|---|----|----|----|
| 1 | 11 | 21 | 31 |
| V | A | G | G |
| D | A | I | Y |
| G | G | G | S |
| R | - | S | A |
| A | F | N | V |
| T | K | N | G |
| V | R | Y | F |
| L | T | A | . |
| . | . | . | . |

2. C-terminal amino acid sequence:

[0030] The present enzyme previously inhibited by DFP was allowed to react with lysyl endopeptidase. The reaction was effected in a 25 mM Tris-HCl buffer, pH 7.5, by using the inhibited SFase and the lysyl endopeptidase at a ratio of 40:1 at a temperature of 24°C for 6 hours. Each of the obtained fragments was subjected to HPLC using a Vydac protein C4 column (4.6 x 25 cm), and separated by linear concentration gradient using 20% through 60% of CH₃CN in 0.1% TFA under various conditions. As a result of the analysis of the fragments, a fragment having no-basic amino acid at a terminal of the peptide was found, and this fragment was identified as a peptide at the C-terminal. The amino acid composition at the C-terminal (corresponding to the amino acids from Glu in the 177 position to Tyr in the 187 position of SEQ ID No. 1) is listed in Table 5. Further, a sequence assumed to be present within the SFase was also obtained (listed in Table 6; corresponding to the amino acids from Tyr in the 68 position to Arg in the 75 position of SEQ ID No. 1).

Table 5:
C-terminal sequence:

| Step | Amino Acid | pmol | Step | Amino Acid | pmol |
|------|------------|--------|------|------------|--------|
| 1 | Glu | 760.5 | 7 | Gly | 839.0 |
| 2 | Ala | 1176.9 | 8 | Val | 1047.5 |
| 3 | Leu | 1318.5 | 9 | Asn | 603.4 |
| 4 | Ser | 353.0 | 10 | Val | 955.5 |
| 5 | Ala | 1126.6 | 11 | Tyr | 951.4 |
| 6 | Tyr | 934.0 | | | |

1 11
...EALSAYGVNVY

Table 6:
Sequence of SFase (unknown region):

| Step | Amino Acid | pmol | Step | Amino Acid | pmol |
|------|------------|--------|------|------------|--------|
| 1 | Tyr | 1843.5 | 6 | Asn | 1460.9 |
| 2 | Thr | 2375.8 | 7 | Val | 2003.5 |
| 3 | Thr | 2266.2 | 8 | Asp | 1896.5 |
| 4 | Thr | 2713.1 | 9 | Gly | 1047.4 |
| 5 | Thr | 2921.1 | 10 | Arg | 620.4 |

1 10
...YTTTTNVDGR...

V. Determination of a DNA sequence for encoding the SFase:

[0031] The terms used in the description of the present invention are defined as follows:

[0032] An "oligonucleotide" indicates a short single-stranded DNA, and can be chemically synthesized by a known method. The oligonucleotide used in the present invention is chemically synthesized through purification by gel chromatography using a Sephadex G50 column and high-performance liquid chromatography (HPLC) using a reverse phase silica gel column, if not mentioned otherwise.

[0033] "PCR" is an abbreviation of polymerase chain reaction, which is a method for enzymatically amplifying a

desired region in DNA (Saiki et al., Science, 239, 487-497 (1988)). First, DNA to be amplified is thermally denatured into a single-strand, and the region at the 3' terminal of each of the obtained single-stranded template DNA is annealed with complementary oligonucleotide primers (two types, one of which is a sense strand and the other an antisense strand). Then, using the action of DNA polymerase, DNA chains are extended from the respective primers. By repeating

5 a series of these reactions, the objective DNA can be amplified a hundred thousand through a million times.
[0034] A "Southern analysis" indicates a method for determining whether or not an objective gene is included in a DNA fragment obtained through cleavage by a restriction enzyme. For Southern analysis, DNA is first digested with a restriction enzyme that recognizes a specific base sequence in the two-stranded DNA for the cleavage thereof. The obtained digest is subjected to 1% agarose gel electrophoresis, then denatured through an alkali treatment into a
 10 single-strand, and the resultant is placed on a nylon filter to be transferred. An oligonucleotide or a DNA fragment, which is part of the objective gene, is prepared and labeled to be used as a probe. The probe is subjected to hybridization with the single-stranded DNA on the nylon filter, thereby analyzing the DNA.

[0035] Next, a method for determining a DNA sequence for encoding the SFase of the invention will be described in the order of procedures. The DNA sequence was determined by analyzing a genome DNA of *Streptomyces fradiae*
 15 ATCC 14544 strain through a method combining PCR, Southern analysis, a direct sequencing method and the like.

(1) Analysis of an internal sequence of a genome DNA by PCR:

[0036] A sequence for encoding the SFase can be obtained, for example, from a genome DNA. For this purpose, a
 20 genome DNA of *Streptomyces fradiae* ATCC 14544 strain is first prepared from cultured cells of the strain by a known method (J. Marmur et al., J. Mol. Biol. (1961) 3, 208-218). The obtained genome DNA is used as a template DNA for PCR. An oligonucleotide primer used in PCR can be synthesized by a general method based on the amino acid sequence around the N-terminal of the purified enzyme obtained in the above-mentioned item IV (6); and the amino acid
 25 sequence of a peptide obtained by partially decomposing the enzyme. For example, an oligonucleotide for encoding the amino acid sequence in the 1 position through 8 position on the N-terminal side of the SFase, i.e., Val Ala Gly Gly Asp Ala Ile Tyr, (up to the second base of a triplet for encoding Tyr at the C-terminal; 23-mer) is chemically synthesized to be used as a sense primer SF1. Since *Streptomyces fradiae* is known to be GC rich, a primer (SEQ ID No. 2) including only one kind of an oligonucleotide is selected to be used without considering degeneracy.

[0037] Separately from the above, a peptide having the highest reliability is selected among the peptides that are
 30 obtained by partially decomposing the purified enzyme with lysyl endopeptidase and sequenced, and an oligonucleotide is synthesized based on the selected peptide. As is described in Example 2 below, since a sequence of Tyr Thr Thr Thr Thr Asn Val Asp is obtained, a 24-mer complementary to the oligonucleotide for encoding this amino acid sequence is chemically synthesized to be used as an antisense primer SF2 (SEQ ID No. 3).

[0038] By conducting PCR using template DNA, the sense primer SF1 and the antisense primer SF2, the objective
 35 DNA chain in the genome DNA is extended and amplified. The obtained PCR product is analyzed through agarose gel electrophoresis, thereby obtaining a DNA fragment having a length of approximately 220 bp. This DNA fragment is introduced into an appropriate vector to conduct subcloning, and then the DNA sequence is determined by a chain terminator method (Proc. Natl. Acad. Sci. USA Vol. 74, 5463-5467 (1977)). As a result, a DNA sequence of 225 bp is determined, and it is found that the amino acid sequence Tyr Thr Thr Thr Thr Asn Val Asp, which is used as a base
 40 for producing the antisense primer SF2, is an amino acid sequence positioned in the 68 position through 75 position.

(2) Southern analysis of the genome DNA:

[0039] The genome DNA derived from *Streptomyces fradiae* ATCC 14544 strain prepared in the above-mentioned
 45 item (1) is double-digested with restriction enzymes Sall and PstI, and the resultant is subjected to agarose gel electrophoresis for separation. Then, the obtained DNA fragments are blotted on a nylon membrane filter for Southern analysis. As a probe for hybridization, the PCR product of the SF1-SF2 of item (1), which was previously labeled with ³²P-dCTP by a general method, is used. Positive DNA fragments hybridized with this probe are exhibited as a
 50 band of 2.0 kb.

(3) Determination of a base sequence of the genome DNA of the SFase:

[0040] The genome DNA of *Streptomyces fradiae* ATCC 14544 strain obtained in item (1) is cleaved with an appropriate
 55 restriction enzyme, the resultant is introduced into an appropriate vector, and the DNA base sequence of the SFase is determined by the chain terminator method mentioned in item (1).

[0041] For example, the genome DNA of *Streptomyces fradiae* ATCC 14544 strain obtained in item (1) is digested with Sall and PstI. The DNA fragment of approximately 2 kb confirmed in item (2) is subcloned with M13mp10 previously treated with the same restriction enzymes, and a plaque hybridization is conducted. At this point, as a probe, the probe

for hybridization used in the Southern analysis of item (2) can be used. Plasmid DNA is isolated from the plaque which is found to be positive through plaque hybridization, and a DNA sequence of the inserted fragment of this plasmid is determined. As a result of the DNA sequencing, the DNA sequence of 2064 bp shown in SEQ ID No. 1 is determined, and SFase is found to comprise a DNA sequence for encoding a prepropeptide comprising amino acids of 170 residues including methionine encoded by a translation initiation codon ATG and a mature peptide comprising amino acids of 187 residues. Upstream, by eight bases, from ATG in the -170 position included in the 5' non-translation region, a sequence AAGGAG of a ribosome binding region is present. Further upstream, a -10 region and a -35 region both including a promoter-like sequence are found. On the 3' side of the DNA sequence for encoding methionine, a basic amino acid specific to a signal peptide is present, to which a hydrophobic amino acid is subsequent. Further, Ala-X-Ala (wherein X indicates an optional amino acid), which is a signal peptide recognition sequence, is also present. Thus, it is found that a typical signal peptide is constructed.

[0042] I. Svendsen et al. (described above) reported an amino acid sequence, of the same type, of an enzyme derived from *Streptomyces griseus*. When the amino acid sequence of this enzyme is compared with that of the SFase of the invention, the homology is found to be 81.3%. Figure 1 shows the comparison between the amino acid sequences of the mature peptides of these two enzymes. The letters used in Figure 1 indicate the ordinary amino acids each with one letter. I. Svendsen et al., however, did not perform gene cloning, and hence, the precursor, the promoter sequence and the like are unknown.

[0043] The present invention provides a novel protease for specifically cleaving the C-terminal of a glutamic residue in the amino acid sequence of a polypeptide and a DNA sequence for encoding the protease. This protease can be prepared by culturing *Streptomyces fradiae* or by chemical synthesis. Such a protease can be used for a variety of purposes including analysis of a protein and cleavage of a peptide chain at a desired site in a fusion protein. The DNA sequence for encoding the protease can be used to produce and detect the protease.

(Examples)

[0044] The present invention will now be described by way of examples.

[Example 1]

(Harvest of the SFase)

[0045] *Streptomyces fradiae* ATCC 14544 strain was cultured in a medium having pH 7.0 and including 3% potato starch, 1% soybean meal, 0.5% corn steep liquor, 0.5% glycerol, 0.3% NaCl, 0.35% CaCO₃ and 0.1% yeast extract at a temperature of 28°C for 3 through 5 days (K. Morihara et al., Biochim. Biophys. Acta., 139 (1967) 382-397). The obtained culture solution was centrifuged (at 8000 rpm for 30 minutes) to obtain a supernatant by removing the fungi. 60% Saturated ammonium sulfate was added to the obtained supernatant, thereby precipitating a protein. The thus obtained precipitate was centrifuged (at 8000 rpm for 30 minutes), and the collected precipitate was dissolved in 10 mM Tris-HCl, pH 7.5, including 2 mM CaCl₂. Then, the obtained mixed solution was dialyzed using the same buffer solution as an outer solution. Next, the dialysis inner solution was lyophilized to obtain crude enzyme powder. Then, 104.4 g of the crude enzyme powder was suspended in 1.2 L of 5 mM Tris-HCl, pH 7.5, and the resultant was dialyzed again by using 20 L of the same type of the buffer solution as an outer solution for 20 hours. The obtained dialysis inner solution was centrifuged (at 8000 rpm for 20 minutes), thereby obtaining a solution of 1.3 L.

[0046] To the obtained solution, 1.0 L of S-Sepharose (equilibrated with 5 mM Tris-HCl, pH 7.5) was added, and the obtained mixture was gently stirred at a temperature of 4°C for 5 hours. Then, the supernatant was discarded, the S-Sepharose having adsorbed the enzyme was charged in a glass column (11.2 x 10 cm) to be washed with approximately 3 L of 5 mM Tris-HCl, pH 7.5, and eluted with 5 mM Tris-HCl, 0.3 M NaCl, pH 7.5.

[0047] Fractions having SFase activity were eluted from the column, collected (2.3 L), and dialyzed with 5 mM Tris-HCl, pH 7.5, for two days and nights (during which an outer solution of 20 L was exchanged three times). Approximately 2.5 L of the dialysis inner material was allowed to adsorb on approximately 500 ml of S-Sepharose (equilibrated with 5 mM Tris-HCl, pH 7.5) charged in a column (5 x 25 cm). Then, the column was washed with approximately 2.5 L of the same type of buffer solution as that used in the equilibrium, and eluted with 5 L of a buffer solution, including 0 through 0.1 M NaCl, by linear concentration gradient. Then, fractions having SFase activity were collected (eluted at a concentration of approximately 0.05 M NaCl: approximately 1000 ml), and dialyzed overnight with 5 mM Tris-HCl, pH 7.5. Next, the dialysis inner material was subjected to affinity chromatography. In the affinity chromatography, approximately 70 ml of CH Sepharose 4B-Phe-Leu-D-Glu-OMe was used as a carrier, which was charged in a column of 2.5 x 14 cm and equilibrated with 5 mM Tris-HCl, pH 7.5. The dialysis inner material was allowed to be adsorbed upon the column. The column was washed with approximately 500 ml of the same type of buffer solution as that used in the equilibrium, and eluted with 1.0 L of the same type of buffer solution, including 0 through 0.3 M NaCl, by linear

concentration gradient.

[0048] The SFase activity of the collected fractions was measured in accordance with the aforementioned activity measuring method. The SFase activity was found in fractions eluted around the NaCl concentration of 0.25 M. The thus obtained fractions having SFase activity were collected (0.2 L), and dialyzed for 24 hours using 5 mM Tris-HCl, pH 7.5, as an outer solution (20 L x 2). Approximately 320 ml of S-Sepharose charged in a column of 3.2 x 40 cm was equilibrated with 5 mM Tris-HCl, pH 7.5, and the obtained dialysis inner material was allowed to be adsorbed upon the column. Next, the column was washed with approximately 1.5 L of the same type of buffer solution as that used in the equilibrium, and eluted with 4 L of buffer solution, including 0 through 0.1 M NaCl, by linear concentration gradient.

[0049] The SFase activity of the collected fractions was measured in accordance with the aforementioned activity measuring method. Further, the absorbance of each fraction at 280 nm was measured to be used as a criterion for a protein concentration. Both results are shown in Figure 2. As is apparent from Figure 2, the SFase activity was found in fractions eluted at a NaCl concentration of approximately 0.05 M (200 ml). The obtained fractions were dialyzed for 40 hours using 5 mM Tris-HCl, pH 7.5 (20 L each was exchanged three times) as an outer solution to be used as a preserved solution. The thus obtained SFase exhibited a single band in SDS-PAGE. From 104.4 g of the crude enzyme powder, approximately 3.0 mg of the enzyme was obtained (quantified using a Bio Rad Protein assay kit).

[Example 2]

(Sequencing of DNA for encoding the SFase)

(1) Analysis of an internal sequence of a genome DNA by PCR:

[0050] A genome DNA was prepared from cultured cells of *Streptomyces fradiae* ATCC 14544 strain by the method of J. Marmur (described above), and the obtained DNA was used as a template DNA for PCR. An oligonucleotide primer to be used in PCR was produced based on a determined amino acid sequence of part of the SFase produced from *Streptomyces fradiae* ATCC 14544 strain. First, an oligonucleotide for encoding the amino acid sequence in the 1 position through 8 position on the N-terminal side of the SFase, i.e., Val Ala Gly Gly Asp Ala Ile Tyr, (up to the second base of a triplet for encoding Tyr at the C-terminal; 23-mer) was chemically synthesized to be used as a sense primer SF1. Since *Streptomyces fradiae* is known to be GC rich, a primer (SEQ ID No. 2) including only one kind of oligonucleotide was selected to be used without considering degeneracy.

[0051] Next, the purified SFase obtained in Example 1 (after DEP treatment) was allowed to react with lysyl endopeptidase (manufactured by Wako Pure Chemical Industries, Ltd.) at a temperature of 37°C for 5 hours. The resultant digest with the enzyme was decomposed and purified by high-performance liquid chromatography using a Vydac protein C4, 300A column (4.6 x 25 cm). The amino acid sequences of the thus obtained digested fragments were checked with a 477 type protein sequencer manufactured by Applied Biosystems. The amino acid sequence of one of the fragments was Tyr Thr Thr Thr Thr Asn Val Asp. Then, a 24-mer complementary to the oligonucleotide for encoding this amino acid was chemically synthesized to be used as an antisense primer SF2 (SEQ ID No. 3).

[0052] By using the template DNA, the sense primer SF1 and the antisense primer SF2, the DNA was amplified by the PCR method [Saiki et al., Science 239, 487-491 (1989)]. When part of the obtained PCR product was analyzed by 1% agarose gel electrophoresis, a DNA fragment of approximately 220 bp was found. This DNA fragment was isolated, blunted through a treatment by a Klenow fragment, and subcloned with M13mp10 previously treated with SmaI. Then, the DNA sequence was determined by the chain terminator method (Sanger et al.; described above). As a result, a DNA sequence of 225 bp was determined, and the amino acid sequence corresponding to the antisense primer SF2 was found to be positioned in the 68 position through the 75 position.

(2) Southern analysis of the genome DNA:

[0053] The genome DNA derived from *Streptomyces fradiae* ATCC 14544 strain prepared in item (1) was double-digested with the restriction enzymes Sall and PstI, and the resultant was subjected to 1% agarose gel electrophoresis. Then, the obtained DNA fragments were blotted on a nylon membrane filter, and Southern analysis was conducted. As a probe for hybridization, the PCR product of the SF1-SF2 of item (1) corresponding to the amino acid sequence from the 38 position to the 45 position of the SFase (SEQ ID No. 4; a probe SF3) was labeled with ³²P-dCTP by a general method. Positive DNA fragments hybridized with this probe were exhibited as a band of 2.0 kb.

(3) Determination of a base sequence of the genome DNA of the SFase:

[0054] The DNA fragments of approximately 2 kb obtained by digesting, with Sall and PstI, the genome DNA of *Streptomyces fradiae* ATCC 14544 strain obtained in item (1) were treated with the same type of restriction enzymes,

and then ligated to dephosphorylated M13mp10. The ligation was performed by using a commercially available kit (manufactured by Takara Shuzo Co., Ltd.).

[0055] Next, plaque hybridization was conducted. At this point, the probe SF3 used in item (2) was used as a probe. Plasmid DNA was isolated from the plaque which was found to be positive through plaque hybridization, and the DNA sequence of the inserted fragment of this plasmid was determined by the aforementioned chain terminator method. As a result of the DNA sequencing, the DNA sequence of 2064 bp was determined. The sequence is shown in SEQ ID No. 1.

[0056] Through study of the DNA sequence of the SFase determined in the above-mentioned manner, it was found to comprise a DNA sequence for encoding a prepropeptide comprising amino acids of 170 residues including methionine encoded by a translation initiation codon ATG and a mature peptide comprising amino acids of 187 residues. Upstream, by eight bases, from ATG in the -170 position included in the 5' non-translation region, a sequence AAGGAG of a ribosome binding region was present. Further upstream, a -10 region and a -35 region both including a promoter-like sequence were found. On the 3' side of the DNA sequence for encoding the methionine, a basic amino acid specific to a signal peptide was present, to which a hydrophobic amino acid was subsequent. Further, Ala-X-Ala (wherein X indicates an optional amino acid), which is a signal peptide recognition sequence, was also present. Thus, it was found that a typical signal peptide was constructed.

(Sequence Listing)

[0057]

SEQ ID No. 1

Length: 2064

Type: nucleic acid

Strandedness: double

Topology: linear

Molecule type: Genomic DNA

Original Source:

Organism: Streptomyces fradiae

Feature:

Name/key: -35 signal

Location: 359..364

Identification method: by experiment

Name/key: -10 signal

Location: 378..383

Identification method: by experiment

Name/key: CDS

Location: 435..1505

Identification method: by experiment

Name/key: sig peptide

Location: 435..944

Identification method: by experiment

Sequence description:

EP 0 647 710 B1

5 GTCGACCGGC GCCGTGGTCG GCCAGCAGCC TTCGGCGGCG GCCGCGTCCG GCACCAACGA 60
 CAAGGCCGGC GCGCCGCAGA ACCTGATGCG CTGGACGCTG ACCCGCGCCA TCAAGGAGAC 120
 GCTGGTCCCG CCGACCGGGT ACGGCTACCC GCACATGCTG ACGCCCGTGC AGCTGACGCC 180
 CGCCCCGGCC CGTACGCACG GCGCGGCGGG CCGACGCCCC GCGCCGCGGC ACCGCGGCGG 240
 GGCTGCGGCC CGCGAGGCCG CCCCCTGACG CCGCCCCCTT CCGGCCCCCA CCGCGGGAGG 300
 GGGCGGTCTC CATGTCCGGG TCGTCGCGCC GGAGGCCGGG CGCGGCGGAA ACGCGCCCTT 360
 GCGGTCAGTG AAGCGTTCAI GTCAAGTTGC TCCCTCGSTA CGTCCCCACC CGTCACCGAC 420
 10 AAGGAGACCC CGGC ATG AGA CGC ACC ACC CGC GCG CGC ACC GGT CTG TCC 470

Met Arg Arg Thr Thr Arg Ala Arg Thr Gly Leu Ser
 -170 -165 -160

15

20

25

30

35

40

45

50

55

| | | |
|----|---|-----|
| | GCC CTG CTC CTC GCC GCC TCC CTC GGC CTC GGC GCC GCC CCG GCC GGA | 518 |
| | Ala Leu Leu Leu Ala Ala Ser Leu Gly Leu Gly Ala Ala Pro Ala Gly | |
| 5 | -155 -150 -145 | |
| | GCC GAC GCC CCC CAG AGG CCC GCC CCC ACC CCG GCC TCC GAC TCC GCC | 566 |
| | Ala Asp Ala Pro Gln Arg Pro Ala Pro Thr Pro Ala Ser Asp Ser Ala | |
| | -140 -135 -130 | |
| 10 | GCC GCC CTC CAC GCC CTC GAC GCC GCC GTC GAG CCG ACC CTC GGC GAC | 614 |
| | Ala Ala Leu His Ala Leu Asp Ala Ala Val Glu Arg Thr Leu Gly Asp | |
| | -125 -120 -115 | |
| 15 | JAC AGC GCC GGC ACC TAC GTG GAC GCC GGC ACC GGC GAA CTC GTC GTC | 662 |
| | Asp Ser Ala Gly Thr Tyr Val Asp Ala Gly Thr Gly Glu Leu Val Val | |
| | -110 -105 -100 -95 | |
| | ACC GTC ACC ACC GAG GCC GCC GCC GCC AAG GTC CGC GCG GCG GGC GCC | 710 |
| | Thr Val Thr Thr Glu Ala Ala Ala Ala Lys Val Arg Ala Ala Gly Ala | |
| 20 | -90 -85 -80 | |
| | ACG CCC CGG CGG GTC CAG CGC GGC GCC GCC GAG CTC GAC GCG GCC ATG | 758 |
| | Thr Pro Arg Arg Val Gln Arg Gly Ala Ala Glu Leu Asp Ala Ala Met | |
| | -75 -70 -65 | |
| 25 | GCC GCC CTG GAG GCA CGG GCC AAG ATC CCC GGC ACC TCG TGG GGG CTG | 806 |
| | Ala Ala Leu Glu Ala Arg Ala Lys Ile Pro Gly Thr Ser Trp Gly Leu | |
| | -60 -55 -50 | |
| 30 | GAC CCG CGC ACC AAC CGG ATC GCC GTG GAG GCC GAC TCC TCC GTC TCC | 854 |
| | Asp Pro Arg Thr Asn Arg Ile Ala Val Glu Ala Asp Ser Ser Val Ser | |
| | -45 -40 -35 | |
| | GCC CGC GAC CTG GCC CGG CTC CGC AAG GTC GCC GCC TCC CTC GAC GGC | 902 |
| 35 | Ala Arg Asp Leu Ala Arg Leu Arg Lys Val Ala Ala Ser Leu Asp Gly | |
| | -30 -25 -20 -15 | |
| | GCC GTC AGC GTC ACC CGC GTC CCC GCC GTG TTC CAG CGC GAG GTG GCC | 950 |
| | Ala Val Ser Val Thr Arg Val Pro Gly Val Phe Gln Arg Glu Val Ala | |
| 40 | -10 -5 1 | |
| | GGC GGC GAC GCC ATC TAC GGC GGC GGC TCG CGC TGC TCG GCG GCC TTC | 998 |
| | Gly Gly Asp Ala Ile Tyr Gly Gly Gly Ser Arg Cys Ser Ala Ala Phe | |
| | 5 10 15 | |
| 45 | | |
| 50 | | |
| 55 | | |

| | | |
|----|---|------|
| | AAC GTC ACC AAG AAC GGC GTT CGG TAC TTC CTG ACC GCC GGG CAC TGC | 1046 |
| | Asn Val Thr Lys Asn Gly Val Arg Tyr Phe Leu Thr Ala Gly His Cys | |
| 5 | 20 25 30 | |
| | ACC AAC CTC TCG TCC ACC TGG TCG TCC ACC TCC GGC GGC ACG TCC ATC | 1094 |
| | Thr Asn Leu Ser Ser Thr Trp Ser Ser Thr Ser Gly Gly Thr Ser Ile | |
| | 35 40 45 50 | |
| 10 | GGC GTC CGC GAG GGC ACC AGC TTC CCG ACC AAC GAC TAC GGC ATC GTC | 1142 |
| | Gly Val Arg Glu Gly Thr Ser Phe Pro Thr Asn Asp Tyr Gly Ile Val | |
| | 55 60 65 | |
| 15 | CGC TAC ACG ACC ACG ACC AAC GTG GAC GGC CGG GTC AAC CTC TAC AAC | 1190 |
| | Arg Tyr Thr Thr Thr Thr Asn Val Asp Gly Arg Val Asn Leu Tyr Asn | |
| | 70 75 80 | |
| | GGC GGC TAC CAG GAC ATC GCC TCC GCG GCC GAC GCC GTC GTG GGC CAG | 1238 |
| 20 | Gly Gly Tyr Gln Asp Ile Ala Ser Ala Ala Asp Ala Val Val Gly Gln | |
| | 85 90 95 | |
| | GCC ATC AAG AAG AGC GGC TCC ACG ACC AAG GTC ACC AGC GGC ACC GTC | 1286 |
| | Ala Ile Lys Lys Ser Gly Ser Thr Thr Lys Val Thr Ser Gly Thr Val | |
| 25 | 100 105 110 | |
| | AGC GCC GTC AAC GTC ACC GTC AAC TAC AGC GAC GGC CCC GTC TAC GGC | 1334 |
| | Ser Ala Val Asn Val Thr Val Asn Tyr Ser Asp Gly Pro Val Tyr Gly | |
| | 115 120 125 130 | |
| 30 | ATG GTC CGC ACG ACC GCC TGC TCC GCC GGC GGC GAC AGC GGC GGC GCC | 1382 |
| | Met Val Arg Thr Thr Ala Cys Ser Ala Gly Gly Asp Ser Gly Gly Ala | |
| | 135 140 145 | |
| | CAC TTC GCC GGT TCC GTC GCC CTC GGC ATC CAC TCG GGC AGC TCC GGC | 1430 |
| 35 | His Phe Ala Gly Ser Val Ala Leu Gly Ile His Ser Gly Ser Ser Gly | |
| | 150 155 160 | |
| | TGC ACG GGC ACC AAC GGC TCC GCC ATC CAC CAG CCG GTC CGC GAG GCC | 1478 |
| 40 | Cys Thr Gly Thr Asn Gly Ser Ala Ile His Gln Pro Val Arg Glu Ala | |
| | 165 170 175 | |
| | CTG AGC GCC TAT GGC GTG AAC GTG TAC TGACCCGCCCG CGCCCGGCCG | 1525 |
| | Leu Ser Ala Tyr Gly Val Asn Val Tyr | |
| 45 | 180 185 | |
| 50 | | |
| 55 | | |

GGCGGAGTGC CACAGTGGGC GGTATGAGTG ATCGTCTCCA TACCGCCCAC ACCTACGAGC 1585
 TGGCACCGGC CGTGCTCGCC TCGATCCGCG CCCTCCTCCA GGACGCCTTC GAAGGCGACC 1645
 5 TGAGCGAGGA GGACTGGGAG CACGGCCTCG GCGGCGTCCA CGCGTACGCC GAGGACCGCG 1705
 GCGGGCGGAT CGTGGCGCAC GGCAGCGTCG TCATGCGCCG CGTGGTCCAC GCGGGGCGCG 1765
 CCCACCGCGT CGGGTACGTG GAGGCCGTGG GCGTACGGGC GGACCGCCGC CGCCAGGGGC 1825
 TGGGCACGCG GGTGATGGCG GCGCTGGAGG AGGTCTCGA CGGGGCCTAC GACTTCGGCG 1885
 10 CGCTCTCCGC CTCGACGCG GCGGCCCTGC TCTAGCTCGC GCGGCTGGCA CCTGTGGAAC 1945
 GGCCGCATCG AGGCGTACGG CCCGCACGGG GTGGTCCGGA TGGCGGACGA GGAGGACTGC 2005
 ACCTTCCTCG GCCCTCGGCG GCGCGCCCGC TGCCGCGGCC GGAGGCGCGT CCCCTGCAG 2064

SEQ ID No. 2

Length: 23

Type: nucleic acid

Strandedness: single

Topology: linear

Molecule type: other nucleic acid, synthesized DNA

Sequence description:

GTCGCCGGCG GCGACGCCAT CTA

SEQ ID No. 3

Length: 24

Type: nucleic acid

Strandedness: single

Topology: linear

Molecule type: other nucleic acid, synthesized DNA

Sequence description:

GTCCACGTTG GTGGTGGTGG TGTA

SEQ ID No. 4

Length: 24

Type: nucleic acid

Strandedness: single

Topology: linear

Molecule type: other nucleic acid, synthesized DNA

Sequence description:

TCGTCCACCT GGTCGTCCAC CTCC

Claims

1. A polypeptide comprising a mature peptide and a signal peptide, wherein said mature peptide is **characterised by** having protease activity for cleaving a peptide bond including a carboxyl group of a glutamic acid residue in a peptide, and having an amino acid sequence from Val in position 1 to Tyr in position 187 of SEQ ID No. 1, and wherein said signal peptide comprises an amino acid sequence from Met in the -170 position to Glu in the -1 position of SEQ ID No. 1.

2. A precursor protein comprising an amino acid sequence from Met in the -170 position to Tyr in the 187 position of SEQ ID No. 1, wherein a mature form of said precursor protein has a protease activity for cleaving a peptide bond including a carboxyl group of a glutamic acid residue in a peptide.

3. A polypeptide or protein according to claim 1 or 2 derived from Streptomyces fradiae.

4. A polypeptide or protein according to claim 1 or 2 derived from Streptomyces fradiae ATCC 14544 strain.

5. A DNA sequence for encoding any one of the polypeptide or protein of claims 1 to 5.

6. A DNA sequence according to claim 5 comprising a base sequence from A in the 435 position to C in the 1505 position of SEQ ID No. 1.

Patentansprüche

1. Polypeptid, das ein reifes Peptid und ein Signalpeptid umfaßt, wobei das reife Peptid **dadurch gekennzeichnet ist, dass** es eine Proteaseaktivität zur Spaltung einer Peptidbindung, einschließlich einer Carboxylgruppe eines Glutaminsäure-Restes in einem Peptid, aufweist und eine Aminosäuresequenz von Val in Position 1 bis Tyr in Position 187 der SEQ ID Nr. 1 aufweist, und worin das Signalpeptid eine Aminosäuresequenz von Met in Position -170 bis Glu in Position -1 der SEQ ID Nr. 1 umfaßt.

2. Ein Vorläuferprotein, umfassend eine Aminosäuresequenz von Met in Position -170 bis Tyr in Position 187 der SEQ ID Nr. 1, worin eine reife Form des Vorläuferproteins eine Proteaseaktivität zur Spaltung einer Peptidbindung, einschließlich einer Carboxylgruppe eines Glutaminsäure-Restes in einem Peptid, aufweist.

3. Polypeptid oder Protein gemäß Anspruch 1 oder 2, von Streptomyces fradiae abstammend.

4. Polypeptid oder Protein gemäß Anspruch 1 oder 2, von Streptomyces fradiae Stamm ATCC 14533 abstammend.

5. DNA-Sequenz, welche für irgend ein Polypeptid oder Protein gemäß Ansprüche 1 bis 4 kodiert.

6. DNA-Sequenz gemäß Anspruch 5, welche die Basensequenz von A in Position 435 bis C in Position 1505 der SEQ ID Nr. 1 umfaßt.

Revendications

1. Polypeptide comprenant un peptide mature et un peptide signal, dans lequel ledit peptide mature est **caractérisé en ce qu'il** possède une activité de protéase pour cliver une liaison peptidique incluant un groupe carboxyle d'un résidu d'acide glutamique dans un peptide, et **en ce qu'il** possède une séquence d'acides aminés depuis Val dans la position 1 à Tyr dans la position 187 de la SEQ. ID. n° 1, et dans lequel ledit peptide signal comprend une séquence d'acides aminés depuis Met dans la position -170 à Glu dans la position -1 de la SEQ. ID. n° 1.

2. Protéine précurseur comprenant une séquence d'acides aminés depuis Met dans la position -170 à Tyr dans la position 187 de la SEQ. ID. n° 1, dans laquelle une forme mature de ladite protéine précurseur possède une activité de protéase pour cliver une liaison peptidique incluant un groupe carboxyle d'un résidu d'acide glutamique dans un peptide.

3. Polypeptide ou protéine selon la revendication 1 ou 2 provenant de Streptomyces fradiae.

4. Polypeptide ou protéine selon la revendication 1 ou 2, provenant de la souche de Streptomyces fradiae ATCC 14544.

5. Séquence d'ADN pour coder l'un quelconque du peptide ou de la protéine des revendications 1 à 5.

6. Séquence d'ADN selon la revendication 5, comprenant une séquence de base depuis A dans la position 435 à C dans la position 1505 de la SEQ. ID. n° 1.

Fig. 1

```

1
(SFase) VAGGDAIYGG GSRCSAAFNV TKNGVRYFLT AGHCTNLSST WSSTSCGTSI
(SGPE) .L..G.....G.A...V. ....I..AN ..AS...SVV
51
GVREGTSFPT NDYGIVRYTT TTNYDGRVNL YNGGYQDIAS AADAVYQQA1
.....D GSSPA.T.D. ...ST...S. ..N.....
101
KKSOSTTKVT SGTVSAYNVT VNYSDCPVYG MYRTTACSAG GDSGGAKFAG
.....T.....G.....N..G.....
151                                     187
SVALGIHSGS SCCTGTNGSA IHQPVREALS AYGVNYY
.....S..A... ..TK... ..T..L
                                     188

```

Fig. 2

